12-4-2013

The Influence and Deception of Twitter: The Authenticity of the Narrative and Slacktivism in the Australian Electoral Process

Benjamin Waugh
*Edith Cowan University*

Maldini Abdipanah
*Edith Cowan University*

Omid Hashemi
*Edith Cowan University*

Shaquille A. Rahman
*Edith Cowan University*

David M. Cook
*Edith Cowan University*

Follow this and additional works at: [https://ro.ecu.edu.au/isw](https://ro.ecu.edu.au/isw)

Part of the *Computer Sciences Commons*

DOI: 10.4225/75/57a849a9befb7
14th Australian Information Warfare Conference, Edith Cowan University, Perth, Western Australia, 2nd-4th December, 2013
This Conference Proceeding is posted at Research Online.
[https://ro.ecu.edu.au/isw/54](https://ro.ecu.edu.au/isw/54)
THE INFLUENCE AND DECEPTION OF TWITTER: THE AUTHENTICITY OF THE NARRATIVE AND SLACKTIVISM IN THE AUSTRALIAN ELECTORAL PROCESS

Benjamin Waugh¹, Maldini Abdipanah², Omid Hashemi³, Shaquille Abdul Rahman³, and David M. Cook¹²
School of Computer and Security Science¹, Security Research Institute²
Edith Cowan University, Perth, Australia
bwaugh0@our.ecu.edu.au, d.cook@ecu.edu.au, mabdipan@our.ecu.edu.au, ohashemi@our.ecu.edu.au, shaquila@our.ecu.edu.au

Abstract
It is uncertain how many discreet users occupy the social media community. Fake tweets, sock puppets, force-multipliers and botnets have become embedded within the fabric of new media in sufficient numbers that social media support by means of quantity is no longer a reliable metric for determining authority and influence within openly expressed issues and causes. Election campaigns, and their associated political agendas, can now be influenced by non-specific virtual presences that cajole and redirect opinions without declaring identity or allegiance. In the lead up to the 2013 Australian Federal Election, the open source Twitter activity for the two major party leaders was examined in order to establish patterns of information diffusion. The results showed fake online personas, fake bots deploying automated Twitter dissemination, and deceptive Twitter strategies. New media tolerates slacktivism, where Twitter users mistake auto-narrative for genuine political sentiment. This study demonstrates the need to increase legitimacy and validity in micro-blogging forms of new media.

Keywords
Twitter, sock puppets, fake personas, botnets, social media, slacktivism, meat-puppets, fake micro-blogging, electoral manipulation

INTRODUCTION
Background to Twitter
Twitter is an online social media micro-blogging service where over 500 million users send and read short messages of 140 characters or less called "tweets" (Lloyd, 2012; Twitter, 2013). The central influence of Twitter draws from two key components; firstly the practice of "re-tweeting" other people’s tweets, and secondly the ability to overtly “follow” other users, providing immediate access to preferred content and confirming interest in both the original tweeter and the associated persons’ narrative (Chu, Gianvecchio, Wang, & Jajodia, 2010; Kwak, Lee, Park & Moon, 2010). Retweets hold considerably more authority and leverage than original tweets because they transmit an original message from a source, whilst at the same time conferring tacit approval of each subsequent retweeted message (Parmelee & Bichard, 2013).

The Influence of Twitter and the Legitimacy of the retweet
Twitter messages are informational narratives that stimulate mass audiences who follow, retweet, or read narratives (Jansen, Zhang, Sobel, & Chowdury, 2009). In the case of international political figures, the content of such narratives command significant authority and influence (Papacharissi, 2010; Cogburn & Espinoza-Vasquez, 2011). Elections in Iran, Egypt and Nigeria have all experienced noteworthy influence from the use of Twitter (Solow-Niederman, 2010; Hamdy, 2010; Howard, 2011). From a national security perspective the ability to stimulate large numbers of people according to political will represents considerable threat (Papacharissi & Oliviera, 2012; Howard, 2011). The legitimacy of that power and influence is therefore of interest, especially if the ability to extend and augment the range and frequency of these messages can be interfered with. It is possible to intercept another’s Twitter account and to post non-genuine messages (Jeffries, 2010), however these rogue messages are easily exposed and easily repudiated. Subsequent messaging can reinforce an individual’s true intentions (Wilson, 2011). Whilst the authenticity of original tweets may be relied upon over the course of time, the validity and accuracy of mass, multiple retweeted messages are less scrutinised (Jewitt, 2009; Papacharissi, 2010). Twitter estimated that at the end of
September 2013 there were approximately 10.75 million fake users (D’Yonfro, 2013) in the form of fake accounts, or accounts belonging to people with multiple personae (USSEC, 2013; Yarow, 2013). Given the ability for Twitter to influence the outcomes of electoral decision making, the need to understand identity authenticity becomes greatly intensified (Lloyd, 2012; Parmelee & Bichard, 2013).

**Fake Tweets: Malfeasance or Meaninglessness**

Opinion is divided as to whether the use of multiple personae represents a legal issue or the right to free speech (Parmelee & Bichard, 2013). Sock puppetry has been used by individuals for hundreds of years as a tactic to deceive and manipulate the thoughts and actions of others (Chu, et al., 2010). The term was first described in online terms in 1993 when an online chatroom discovered one of their participants attempting to sway the conversation with a second online persona under a false pseudonym (Rollins, 1993). Since that time, the terminology and practice of sock puppetry have been applied to online marketing (Streitfield, 2012), political support (Cogburn & Espinoza-Vasquez, 2011), and terrorist coercion (Conway, 2012).

In the case of elections, Twitter is a powerful vehicle for persuasion, assisting in nations where the normal means of media becomes blocked for those in opposition to government (Bartlett, Birdwell, and Littler, 2011) and assisting in democratic nations where minority voices seek to get their messages heard (Gleason, 2013). The ability to deploy force multipliers through Twitter traffic means that the retweets of a few can become a significant numerical retweet statistic in support of a candidate, policy or message of significance. The Iranian elections of 2009 demonstrated widespread slacktivism, where Twitter followers mistook blog feeds and botnets for genuine voices of political conviction (Christensen, 2011). In two party dominated elections such as those in the United States and in Australia, where emphasis is placed on support for one political leader over another, party machines may seek to redouble their effort through the use of sock puppetry to augment the perception of support. In the case of the 2013 federal elections in Australia, both the incumbent Prime Minister and the Leader of the Opposition were credited with significant numbers of fake Twitter followers (Butt & Hounslow, 2013).

**Meat-puppets and Botnets: retweets for hire**

The notion that large numbers of on-line followers can be ‘switched on’ demonstrates the increased practice of meat puppetry, where there is an intentional paid-for group-harnessing of Twitter support for a particular person or narrative. Kevin Ashton’s imaginary motivational speaker ‘Santiago Swallow’ was famously elevated in importance by the addition of 90,000 fake followers for the sum of US$50. Ashton created the fake account in less than 2 hours, searched the website fiverr.com for people selling Twitter followers, and created Santiago Swallow on the 13th of April 2013. He was then able to obtain reports from legitimate social media analyst firms such as PeopleBrowsr, who confirmed Santiago Swallow had an @Kred (2013) influence score of 754 out of 1000 (Ashton, 2013, Butt & Hounslow 2013).

The use of online social networks to spread misinformation and propaganda is most clearly seen in Twitter, where web-based botnets command a considerable portion of twitter traffic (Boshmaf, Muslukhov, Beznosov and Ripeanu, 2011). Botnets form an unethical but significant segment of the twitter community (Chen, 2010). Ashton’s Twitter creation was ‘activated’ largely through the use of a trial copy of TweetAdder, using a simple application that automatically dispensed ‘Santiago Swallow’s tweets on a systematic basis (Ashton, 2013). This automated program quickly established a record of Twitter activity, and borrowed popular platitudes from online sources to establish a persona that appeared current and fashionable.

**Depicting fakes from real users**

Whilst there is substantial evidence of widespread sock puppetry throughout the Twitter community (Krebs, 2011; Wheatley, 2013), it is particularly prominent within election phases, where the support
for a political leader may influence the voting decision of large numbers of people (Parmelee & Bichard, 2013). It therefore becomes important to determine the fake tweets from the real ones. In political terms (assuming prominent leaders have not been denied access to their own Twitter accounts) the measure of influence is best determined in terms of retweeted narrative (Ibid, 2013). Chu et al., (2010) developed a method for determining whether retweets were created by humans, bots or cyborgs by distinguishing retweeting humans as legitimate followers who subscribe to the original authorship of others.

Twitter does not repeatedly challenge retweeting entities to establish whether a bot-like entity is in play. It only asks for a CAPTCHA image during the registration and set up of a Twitter account. As a result, as soon as login has been performed, bots can perform the majority of human twitter activities by calling on Twitter APIs. Sitting between humans and bots there is a third group which Chu et al.,(2010) call cyborgs. These Twitter entities are part-human and part-bot (Edwards, Edwards, Spence and Shelton, 2013). In some instances they are human-assisted bots and in other instances they are bot-assisted humans. One example might be as follows: a human Twitter subscriber logs in and sets up a number of automated feeder programs such as RSS feeds and Blog widgets. The Twitter entity then proceeds to retweet a number of messages showing regular activity. The subscriber then also revisits his cyborg entity and further enhances its ‘humanity’ by posting the occasional message to interact with other known friends. The resultant Twitter traffic looks decidedly human, even though the great majority of the traffic is automated.

In order to depict automated Twitter entities (whether as bots or cyborgs), Chu et al (2010) proposed a four way test to distinguish fakes from humans. Measuring the intervals between retweets proved a very reliable method of sensing automated messaging. Looking for spam was also reliable because humans very rarely message spam, and examining the account properties of each subscriber also proved reliable, since those subscribers with no real account details, pictures, or descriptors, were very rarely indicative of individual humans. Additionally, bots are far more likely to post URLs than humans. By looking at all of these variables in concert, the credibility of retweeting followers can be established (Chu, et al., 2010).

| 1. | Entropy Test - Measure Retweet intervals |
| 2. | Spam and Miscreant Test - Check for Benign or Malicious content |
| 3. | Account Properties – Does the Account have subscriber details or does it look hollow |
| 4. | Discrimination Analysis – combining Entropy, Spam, and Account Properties to evaluate all three indicators |

*Table 1. Chu et al. (2010) 4 way test to distinguish humans from bots and cyborgs*

**Political Twitter Deception**

Twitter has the potential for increasing political participation (Hamdy, 2010; Wilson, 2011; Parmelee & Bichard, 2013). There is a close association with augmented retweets in political elections that suggests the presence of wide-spread sock puppetry (Stieglitz and Dang-Xuan, 2012). Sock puppetry in Twitter retweets is assisted by the use of metadata, in particular URLs and #hashtags (Suh, Lichan, Pirolli, & Chi, 2010; Yang & Counts, 2010). Retweeted integration includes the use of policy-event ‘#hashtags’, to gauge the popularity or significance of events, policies, or people. Thus in political Twitter communities, where a specific event such as an election is in play, the retweet components allow for the amplification of political narratives through sock puppetry.

The transition from sock puppets to meat-puppets marks the elevation from subscribers with a handful of fake personas to the business of invoking thousands of automated fake followers who are paid for, ready, and willing to retweet upon command (Wheatley, 2013). Meat puppets are
essentially ‘guns for hire’ able to be marshalled at a moment’s notice. The Russian election in 2011 drew the attention of fraudsters who harnessed 25000 Twitter accounts in order to send 440,000 retweeted messages disrupted the natural discourse of narrative and counter narrative centred on the election and its associated issues (Thomas, Grier, and Paxson, 2012). Subsequent investigation revealed a spam as a service botnet that controlled over 975,000 Twitter accounts and mail.Ru email addresses (Thomas et al, 2012; Krebs, 2011). The majority of the IP addresses associated with the disruption originated from outside Russia, further demonstrating the global ability for sock puppetry to influence domestic politics (Krebs, 2011).

Businesses offering meat-puppets are overt in their offerings and direct in their dealings. The British firm Buy More Followers advertises a range of new media services with scalable packages of up to 100,000 followers who have the appearance of authenticity (Evon, 2013). Other online entities that trade in increased automated Twitter followings and spam-as-a-service retweets include Fast Followerz, Deyumi, and FollowerSale (TwitterTop, 2013). Global traders such as these offer a range of influence services that span beyond Twitter retweets into spamming, scamming, phishing, and maliciously infiltrating the overall balance of Twitter traffic (Wheatley, 2013). The effect of sock puppetry on this level is above the status of ‘nuisance’ and indicates the value of buying support to sway the perceptions and views of Twitter users to influence the outcomes of national and international elections.

The tolerance of large numbers of noticeably non-genuine Twitter followings is described as ‘slacktivism’, where active Twitter users operate alongside fake Twitter entities yet accept the practice in its multiple forms (Christensen, 2011). Slacktivists increasingly put up with a range of new media activity that incorporates augmented and distorted postings, narratives, and social participation (Rotman, Vieweg, Yardi, Chi, Preece, Shneiderman, Pirolli and Glaisyer, 2011). Low cost, low risk, technology-mediated participation has created the opportunity for negative outcomes expressed through socially driven force multipliers. Slacktivism reduces the the capability for meaningful interpretation of activism when that participation operates alongside fake microblogging in the form of botnets and cyborgs that are either overlooked or unseen (Christensen, 2011).

**Case Study into the Retweet data during the 2013 Australian Federal Election**

The 2013 Australian federal election catalyzed reactions to the retweet activity that followed the two main party leaders vying for Prime Minister. Tony Abbott and Kevin Rudd both had significant followings on Twitter. Using Abbott and Rudd’s leadership tweets, this case study examined 30,535 first generation retweets with a total of 10,201 retweets from Abbott tweets and 20,334 retweets from Rudd tweets. Under similar electoral conditions Chu et al., (2010) and Wilson (2011) predicted significant impact from fake retweets, bots and cyborgs. With the election on the 7th of September 2013, the Twitter activity in the immediate pre- and post- election weeks allowed an examination of retweeter behavior. This study hypothesizes that it is possible to distinguish fake and automated retweets by comparing retweets with key election dates that generated high-interest policy announcements as well as the followers’ inactivity before the election, and their inactivity after the election. A second hypothesis posits that slacktivism guides the Twitter behavior of politically interested followers. This hypothesis looks to demonstrate the emergence of ‘social –loafing’ where followers will retweet political and election narrative without regard for author identity, or narrative information. The study aimed to show larger than normal numbers of Twitter followers, botnets, sock puppets, and their associated retweets in support of political leaders.

**Methodology**

In order to examine online fake electoral personas in Australia, the Twitter accounts (@TonyAbbottMHR and @KRuddMP) of the two main leaders vying for the position of Australian Prime Minister were used as starting points to measure the associated followers, and their
The samples were collected from open source content, available by following the accounts through Twitter.com. Building on Chu’s (et al., 2010) original test using four criteria, a nine-way test was derived that included the original four of Chu’s variables, and added markers for specific event and date-based activity (Table 2.). The test also looked for accounts which were characterized by ‘exclusively single generation’ retweet behavior, on the basis that bots don’t retweet other bots.

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Mass retweets on policy-specific days and times</td>
</tr>
<tr>
<td>7. Follower alignments – Bots don’t follow other Bots</td>
</tr>
<tr>
<td>6. Inactivity after the election</td>
</tr>
<tr>
<td>5. Inactivity before the election</td>
</tr>
<tr>
<td>4. Accounts Created on or about August the 4th 2013. (Announcement of Election)</td>
</tr>
<tr>
<td>3. Account Properties – Does the Account have subscriber details or does it look hollow</td>
</tr>
<tr>
<td>2. Spam and Miscreant Test - Check for Benign or Malicious content</td>
</tr>
<tr>
<td>1. Entropy Test - Measure Retweet intervals</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nine Way Test for Twitter entities in an Election</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Mass retweets on policy-specific days and times</td>
</tr>
<tr>
<td>7. Follower alignments – Bots don’t follow other Bots</td>
</tr>
<tr>
<td>6. Inactivity after the election</td>
</tr>
<tr>
<td>5. Inactivity before the election</td>
</tr>
<tr>
<td>4. Accounts Created on or about August the 4th 2013. (Announcement of Election)</td>
</tr>
<tr>
<td>3. Account Properties – Does the Account have subscriber details or does it look hollow</td>
</tr>
<tr>
<td>2. Spam and Miscreant Test - Check for Benign or Malicious content</td>
</tr>
<tr>
<td>1. Entropy Test - Measure Retweet intervals</td>
</tr>
</tbody>
</table>

Table 2. A nine-way test to distinguish fake retweets in the 2013 Federal Election (adapted from Chu et al, 2010)

Analysis of Results
The key variables showed large numbers of probable bots based upon their retweet intervals. Humans tend to retweet at different times of the day, and in different ways. Bots that are automated to retweet will send narratives at very exacting intervals (eg every 4 hours, or at exactly the same time each day). The increased presence of spam showed probable botnets, since humans rarely retweet spam. In combination, the presence of bots increases exponentially, since humans rarely retweet spam, and if they did, they would be extremely unlikely to retweet that spam at the exact same time each day. Account details are also useful markers. Accounts that show few details, no pictures, and with default background settings are more likely to be bots. Again, in combination with the other variables, the likelihood of certain followers being fake bots increases appreciably.

Other ‘date-specific’ variables offer even more divergence from human followers. Twitter accounts that were activated either on, or just after, the announcement of the federal election date show higher probability of being botnets. Similarly, large numbers of followers of both Tony Abbott and Kevin Rudd showed almost complete inactivity before the election announcement, and after the election result was announced. Those retweets that showed similar followings with each other, shared a positive relationship with other botnet characteristics, but never retweeted each other, look remarkably automated. Bots don’t retweet other bots. Key policy dates also showed strong co-variation between increased retweets and the combined nine-way test variables. Throughout the data sample there existed strong rationale and reasonable explanation suggesting the deployment of bots.
Specific dates also revealed the likely presence of botnets. As previously discussed, large numbers of followers started retweeting after the 4th of August (the announcement of the election date). These included Twitter accounts that had lain dormant for many months, or whose creation occurred on or just after the 4th of August. On August the 11th the numbers of followers for Tony Abbott increased dramatically, suggesting the acquisition of bots on a meat-puppet scale. Abbott’s followers went from 165,235 to total of 234,167 followers, a rise of 68932 within 48 hours (Ralston, 2013). After media conjecture that Abbott had acquired the services of meat-puppets, the numbers dropped by 43,357 followers within the following 24 hour period (Ralston, 2013). The same characteristic aligned positively with Rudd’s Twitter followers. Again, on the 12th of September more than 50000 followers of Kevin Rudd suddenly dropped off the list of Rudd followers after media speculation that the Twitter accounts had been ‘hired’ as a force multiplier to portray support for Kevin Rudd’s election campaign (Andrews, 2013). Whilst it is possible that more than 100000 individual Twitter followers joined up to Rudd and Abbott with 24 hours, in combination with media speculation and the subsequent mass dumping of Twitter accounts, the more likely proposition is that these Twitter accounts are fake, and have been acquired en mass.

The impact of fake Twitter accounts is disguised amongst the combined traffic of genuine human accounts and fake Twitter multipliers. The analysis showed an amalgamation of paid-for botnets, and party-aligned cyborgs (party faithful users with multiple Twitter accounts). The paid-for botnets showed strong positive grouping with each other, often behaving in exactly the same manner, and retweeting exactly the same narratives, and where possible, the same overt metatagging of a URL within the 140 characters of the Twitter message. These retweets show as automatic feeds that are re-expressed so as to imitate individual messages of party support with the aggregated influence of large numbers of retweets. Typical groupings of these retweets present in groups of several thousand at a time, with retweeted messages deploying at regular (robotic) intervals. The cyborg retweets were more difficult to detect, since although they displayed many of the same automated response characteristics of the large group meat-puppets, would also show occasional personalized messages. These personalized messages reinforced the idea that Twitter accounts were real individuals, when in fact they were multiplied expressions of a handful of users. Cyborgs present in smaller groupings, typically from 5-15 expressions. Some of their ‘personalised’ messages were identical to the ‘personalised’ messages of other fellow cyborgs, simply messaged at different (although automated) intervals.

Limitations on botnets and cyborgs
Whilst large numbers of retweets display the obvious characteristics of automation and the presence of fake personas, there is also the possibility of some humans behaving as bots (Gianvecchio, Xie, Wu and Wang, 2008). One of the inherent attractions of Twitter (more than other
social media platforms) is the ability to simply copy and retweet someone else’s narrative (Edwards et al., 2013). Twitter allows for otherwise shy people to engage in an information exchange that builds on the repetition of someone else’s message (Celli, 2011; Zhao and Rosson, 2009). Additionally politics can be a confronting topic to engage with individually and socially (Morozov, 2009). There is a small percentage of the data that has been aggregated as either botnets or cyborgs who are quite possibly human (Edwards et al., 2013).

**Slacktivism in Political Retweets**

There is conjecture about the validity of the support that flows from retweeted political narrative (Christensen, 2011). Slacktivists are deemed as social activists who mistakenly confuse auto-narrative for genuinely supported political posturing. Whilst the media has covered many high profile examples of Twitter deception, the on-line public has remained relatively silent. In the case of the Australian federal election, there would be cause for concern if micro-blogging were seen to deceive the public. Section 329 of the Australian Commonwealth Electoral Act of 1918 states:

Misleading or deceptive publications:

“A person shall not, during the relevant period in relation to an election under this Act, print, publish or distribute, or cause, permit or authorize to be printed, published or distributed, any matter or thing that is likely to mislead or deceive an elector in relation to the casting of a vote.” (Commonwealth Electoral Act, 1918).

The question arises as to whether any electors in the 2013 Australian federal election were sufficiently influenced by the retweeted support or multiplied narrative in terms of their voting behavior. Do retweets constitute ‘published material’? If botnets and cyborg-generated spam represents misleading electoral advertising (by virtue of its existence through fake personas), then there is reason to examine the legal means of distinguishing between real retweets and those from their fake, multiple identities (AEC, 2013). If we assume that botnets are widespread in the same way that tweets are ubiquitous, then we could assume that any and all retweets should be treated as unverified. However, if the retweets of political election narrative were treated in the same way that political messages on printed posters are considered, then the Australian Electoral Commission would need to take action (AEC, 2013). Instead, this paper assumes that slacktivism drives public acceptance of retweeted micro-blogging regardless of frequency or pervasiveness.

**Hypotheses 1 and 2 Results**

The first hypothesis asked whether it was possible to distinguish fake, automated retweets from human ones. The nine-way test revealed more than 28000 followers (combines Rudd and Abbott data) that aligned with the combined features of entropy, spam, identity hollowness, automated retweet intervals, and pre or post retweet inactivity. A combination of Twitter botnets and cyborgs, in the form of both sockpuppets and meat-puppets, interacted regularly with the postings of either the @TonyAbbottMHR or the @KRuddMP Twitter accounts from the period of the 4th of August through until the 18th of September in 2013. The second theory builds upon the positive results of the first hypothesis. Since the detection of large numbers of automated retweets from fake personas operating Twitter accounts was openly discussed in mainstream media, yet not debated beyond initial reporting, a reasonable conclusion is that political parties not only tolerate slacktivism, they accept the benefit of ‘slacker activists’ whether in human, botnet or cyborg form.

**CONCLUSION**

The 2013 Australian Federal Elections were subject to large numbers of automated, non-trustworthy, fake Twitter followers who retweeted messages in support of the two opposing political leaders Kevin Rudd and Tony Abbott. These followers included a combination of botnets, cyborgs, sock puppets and meat-puppets. Their presence in large numbers implies that Twitter support as reported is not a reliable metric for depicting the impact and influence of political issues and their online
discourse. The lack of reaction to obvious sock puppetry and its subsequent withdrawal, implies that slacktivism is an accepted component in new media. Any electoral process that purports to explain support or popularity for an issue or candidate on the basis of Twitter retweets should disclaim the significant distortion that fake personas displace into the political landscape.

REFERENCES

@Kred. (2013) Kred Social Media Influence Platform, retrieved from Twitter on the 23rd September 2013 at: https://twitter.com/Kred


Twitter. (2013). Twitter; An Information Network, retrieved 25th September 2013 from: https://twitter.com/about


